IP Telephony: Architecture and Growth Factor

Tesfa Haile
Jackson State University

Abstract

IP telephony is used to send service like voice, facsimile, and /or voice-message applications between two or more computers in real time. It combines voice and data that are transported via the Internet, rather than the Public Switching Telephone Network (PSTN). IP telephony became a reality because of the H.323 specification. H.323 establishes standards for compression for audio and video data stream.

Introduction

Internet can provide business with a world-wide audience, what about the phone bill you will incur for communication for long distance and international calls? Fortunately, the Internet promises even bigger communication savings by drastically lowering long distance phone bills. Internet Protocol, one of the fastest -growing uses for the Internet, cuts long-distance phone charges down to a bare minimum by using existing data lines and gateways, which translate analogy calls into digital information and back again.

Regular phone calls travel through several circuit switches, while Internet calls need to travel through gateway only twice: once to turn it to digital form and once to change it back. Dedicated lines are not needed. Internet telephony is estimated to grow rapidly. Many analysts estimated that by the year 2010, half of the long-distance traffic will go over IP. Instead of paying between 10 and 30 cents per minutes for domestic long-distance calls, most IP telephony services charge an average of 8 cents per call. Cost savings on international call via the Internet can be even greater.

The major drawback of using IP telephony is that it doesn't give the stable quality that many people expect during a phone call. If the lines are jammed with other Internet traffic, calls can break up and bits of the conversation will get delayed or lost. Since it uses IP/UDP, calls may be unreliable. IP telephony call can be made three ways: computer-to-computer, computer-to-phone, or directly phone-to-phone.

Discussion

Computer Telephone Integration (CTI) provides an environment for application creating that gives access to the function of both telephone system and the computer system. IP telephony is an implementation of CTI. Telephony basically covers the telephone system itself: the way in which telephone networks carry calls, and the way that those telephone calls get to the right person. The Public Switched Telephone Network (PSTN) uses and entire telephone channel for every phone call, fax, or data connections, and route the call from the sender to receiver as if establishing a single end to end circuits. That is why they are called "circuit-switched" networks. Circuits are reserved between the origination switch and the terminating switch based on the called party number (Walter, 1993).

The combination of IP and telephone gives IP telephony. IP telephony sends services like voice, facsimile, and/or voice-message applications between two or more computer users in real time thereby combining voice and data that is transported via the Internet, rather than the PSTN (Siemens). How does IP telephone works? The services starts from the Personal Computer(PC) that has special software to convert the sound into digital codes, which are then passed on to the Internet Service Provider(ISP). The ISP breaks the digital messages up into packets. Packets piece of message each encoded with a destination address. The packets go to the the Internet Telephone Service Provider (ITSP) who reassembles the packets as they arrive and converts them to speech. It then goes to the traditional PSTN, which direct the call to the right phone number. ITSP charges for the local call and a handling fee (Armstrong, Sandler, Elstrom, 97).

H.323 Standard

H.323 establishes standards for compression and decompression of audio and video data streams, ensuring that equipment from different vendors will have area of common support. H. 323 is not tied to any hardware or operating system. It addresses the core Internet-telephony application by defining how delay-sensitive traffic gets priority transport to ensure real-time communications service over the Internet. Video and voice traffic are bandwidth-intensive that could clog the corporate network. H.323 addresses this issue by providing bandwidth management. Network managers can limit the number of simultaneous H.323 connections within the network or the amount of bandwidth available to H.323 applications. These limits enure that critical traffic will not be disrupted. H.323 allows customers to use multimedia applications without changing their network infrastructures since it is designed to compensate for effect of highly variable LAN latency (Siemens).

Customers make conference without worrying about compatibility at the receiving point. The H.323 ensures that the receiver can decompress the information without any problem. H.323 allows customers products to inter operate with other H.323 compliant products by providing device-to device, application -to-application , and vendor-to -vendor interoperability.

Gateway server allows a full-duplex conversion. A gateway consists of several parts. One is a switched-circuit network interface, incorporating T1 of ISDN PRI interface cards. Gateway usually also contains NICs for communication with devices on H.323 network. Other components include digital signal processors, which take care of voice compression and echo cancellation, and a control processor that oversees all other gateway functions (Ong, 1999).

Gatekeepers are required to perform four functions. First, they must translate terminal and gateway LAN aliases to IP or IPX addresses. Second, gatekeepers perform bandwidth control, which involves allocating bandwidth out of the telephone. Gateways have overcome an IP telephone problem of addressing.

Another important piece of any H.323 network is the gatekeeper, which acts as the central point for all calls within its zone. A gatekeeper's zone is defined as H.323 terminals, translation gateways, and multipoint units over which it has control during a call; they can refuse to create more connections once a pre-established upper limit for a number of simultaneous conversations has been reached. A third gatekeeper function is admission control, which use Remote Access Service (RAS) messages to authorize network access. The fourth required function is zone management, which involves performing the previous three tasks for all terminals, gateways, and MUCs within its zone (DataBeam, 1998).

Gatekeepers can also perform several optional functions. One is call-control signaling, which permits the gatekeeper to process Q.931 signal messages. Q.931 is a signal protocol which sets up and terminates a connection between two H.323 devices. A gatekeeper may also perform bandwidth management, an extension of bandwidth control, which means it can determine when there is no available bandwidth for a call, or if there is no more available bandwidth when a call in progress request more. Other optional gatekeeper services include call authorization, which involves the acceptance or rejection of calls based on certain criteria such as time of day, type of services, and lack of bandwidth. Gatekeepers also may perform call management, which involves keeping track of H.323 calls in progress to determine which terminals are busy. This helps gatekeepers redirect calls or save call-setup time by not trying to reach a terminal already in use.

Although H.323 is geared toward audio and video conferencing, it does support data conferencing. ITU incorporates the T-120 standard, which defines point-to-point and multipoint data conferencing sessions with ease. The T-120 data conferencing standard contains a series of communication and application protocols and services that provide support for real-time. This provides exceptional benefits to end user, vendors, and developers tasked with implementing real-time applications. T.120 error-corrected data delivery ensures that all endpoints will receive each data transmission. In multicast enable works, T.120 can employ reliable and unreliable delivery services. Unreliable data delivery is also available without multicast. The T.120 standard supports a broad range of transport options, including PSTN, ISDN, Packet Switched Digital Networks(PSDN), Circuit Switched Digital Network, and popular local area network protocols(DataBeam, 1998).

Architecture Overview

The H.323 Recommendation covers the technical requirements for audio and video communications services in LANs that do not provide a guaranteed Quality of Service (QoS). H.323 references the T.120 specifications for data conferencing and enables conferences which include a data capability. The scope of H.323 does not include the LAN itself or the transport layer that may be used to connect various LANs. Only elements needed for interaction with the Switched Circuit Network are within the scope of H.323. H.323 defines four major components for network-based communications system: Terminal, Gateways, Gatekeepers, and Multipoint Control Units.

Terminals are the client endpoints on the LAN that provides real-time, two-way communications. All terminals must support voice communications, video, and data. H.323

specifies the modes of operation required for different audio, video, and/or data terminals to work together. All H.323 terminals must support H.245, which is used to negotiate channel usage and capabilities.

Gateway is an optional element in an H.323 conference. Gateways provide many services, the most common being a translation function between H.323 conferencing endpoints and other
terminal types. In addition, the Gateway also translates between audio and video codecs and performs call setup and clearing on both the LAN side and switched-circuit network side. In general, the purpose of Gateway is to establish links with analog PSTN terminals, create links with remote H.320-compliant terminals over ISDN-based switched-circuit networks, and establish links with remote H.324- compliant terminals over PSTN network.

Gatekeeper is most important component of an H.323 enabled network. It acts as the central point for all calls within its zone and provides call control services to registered endpoints. In many ways, H.323 gatekeeper acts as a virtual switch.

The Multipoint Control Units (MCU) supports conferences between three or more endpoints. Under H.323, an MCU consists of a multipoint controller, which is required to have one or more multipoint processor. The multipoint processor handles H.245 negotiates between all terminals to determine common capabilities for audio and video processing.

Key Benefits of H.323

H.323 establishes standards for compression and decompression of audio and video data streams, ensuring that equipment from different vendors will have some area of common support.

Interoperability

Users want to conference without worrying about compatibility at the receiving end. Besides ensuring that the receiver can decompress the information, H.323 establishes methods for receiving clients to communicate capabilities to the sender.

Network Independence

H.323 is designed to run on top of common network architectures. As network technology evolves, and as bandwidth-management techniques improves, H.323-based solutions will be able to take advantage of those enhanced capabilities.

Platform and Application Independence

H.323 is not tied up to any hardware or operating system. H.323-compliant platforms will be available in many sizes and shapes. Including video-enabled personal computers, dedicated platforms, IP-enabled telephone handsets, cable TV set-top boxes and turnkey boxes.

Other key benefits of H.323 includes: Multipoint Support, Bandwidth Management, Multicast support, Flexibility and Inter-Network Conferencing.

Quality of Service Challenge in IP Telephony

The Public Switched Telephone Network (PSTN) was designed for real-time voice transport, providing low latency, guaranteed bandwidth, and lossless transmission. On the contrary, IP networks were designed to transport data, they introduces latency, packet jittering, and packet loss.

QoS Challenge: Latency

For toll quality voice conversation, end to end delay must be under 250ms, requiring 100ms gateway delay and less than 150ms IP network delay. IP telephony uses Real Time Protocol (RTP) embedded in IP/UDP. It contains time stamp and sequence number. It has no built-in latency guarantee.

QoS Challenge: Packet Jitter

IP packet arrives time is not deterministic, this creates inter-arrival time jitter. Jitter contributes to latency. Jitter buffer management is needed in gateway.

$$\text{Latency(total)} = \text{Latency(avg)} + \text{Jitter(max)}$$

QoS Challenge: Packet Loss

IP packets may be lost during transmission. It may be discarded by routers. It arrives out-of-order beyond allowable latency. Many gateways have processing to address mitigate packet loss. Some of the techniques used are: packet redundancy, forward error correction, and lost packet interpolation schemes. The goal is to keep packet loss below 5-10%.

Potential QoS Solutions

The following suggestions are provided to improve quality of service: over -provisioning, buy grade of service guarantees, and utilize new IP-related protocols. The over-provisioning allows to build excess capacity in LAN or private WAN and continual capacity planning as needed.

With IP telephony technology being relatively new there are some advantages and disadvantages. Many corporations wanted proof of cost savings, quality and reliability, endorsements by others in the field, and maturity in the technology. The ultimate objective, of course, is reliable, high-quality voice service, the kind that users expect from the PSTN. Internet’s limited bandwidth often results in congestion, which, in turn can cause delays in packet transmission. Packet loss, overhead, latency and jitter are additional disadvantages with IP Telephony technology.

A characteristic that determines the quality of an Internet telephony connection is simply how well the transmitted voice quality matches the speaker's natural voice. If certain voice packets are delayed beyond a specific threshold, Internet telephony software will "guess" at the contents of the lost packets by analyzing the surrounding packets and interpolating. The more interpolation than necessary, the more the voice quality is distorted.

Because the Internet is a packet-switched of connectionless network, the individual packets of each voice signal travel over separate network paths for reassembly in the proper sequence at their ultimate destination. While this makes for a more efficient use of network resources than the circuit-switched PSTN, which routes a call over a single path, it also increases the chances for packet loss. In voice communications, packet loss shows up in the form of gaps or periods of silence in the conversation, leading to a "clipped-speech" effect that is unsatisfactory for most users and unacceptable in business communications.

Packetizing voice codes are becoming better at reducing sensitivity to packet loss. The main approaches are smaller packet sizes, interpolation (algorithmic regeneration of lost sound), and a technique where a low-bit-rate sample of each voice packet is appended to the subsequent packet. Through these techniques, and at some cost of bandwidth efficiency, good sound quality can be maintained even in relatively high packet loss scenarios.

As techniques for reducing sensitivity to packet loss improve, so a new opportunity emerges for the achievement of even greater efficiencies by suppressing the transmission of voice packets whose loss in determined by the encoder to be below a threshold of tolerability at the decoder. This is particularly attractive in the emerging packet-based networking world where random packet loss is being controlled through complex protocols and through server control and where statistical multiplexing favors the reuse of recovered bandwidth.

In IP Telephony each voice packet carries a header with identification and routing information that contributes 'overhead' not present with circuit switching techniques. Real-time traffic (including voice) is very sensitive to delay and packet loss. Transmitting the real-time traffic in very small packets offsets this sensitivity. The use of very small packets has the unfortunate effect of increasing the ratio of overhead to revenue-bearing traffic reducing the effective capacity of a given transmission medium.

A typical encoded voice payload size is 20 to 40 bytes, while overhead in packet-based networks varies between 10 and 40 bytes. Thus, packet overhead consumes and average of 20% to 50% of the occupied capacity when packet networks carry encoded voice traffic. IP networks, in particular the Internet, are among the worst offenders in this area. Evolving methodologies are reducing this overhead through session multiplexing within packets, header compression and other techniques that replace the header with session or local channel identifiers similar to Frame Relay's Data Link Connection Identifier (DLCI). Thus, it is likely that packet-based networks will soon exceed the efficiency of circuit-based networks for the transmission of high-quality voice traffic.

Header overhead in packet-based networks does however provide a significant advantage over its circuit-based networks does however provide a significant advantage over its circuit-based counterpart, delivering the statistical multiplexing capability so valuable for mixing multiple traffic types over a single network (Couch, 1999).

Latency is another challenge to consider with IP Telephony technology. Latency is the time required for data to travel from point A to point B. During a telephone call end-to-end delay much above 300 milliseconds is perceptible. Though many acceptable voice services exceed this level, offerings must operate below it. Advanced packet-based voice solutions keep total delay below the 300ms threshold, but most networks experience higher delay, impeding support for toll-grade voice solutions. This is particularly true for Internet and international traffic. Thus, limiting delay while improving other quality factors is the focus of much of the current research and development in packet-voice technology.

With the advent of IPV6 (RSVP) and Winsock 2.0, many of the issues associated with this delay problem(namely routing, topology, traffic profile and protocol factors) are significantly reduced. .

Jitter is the variability in latency. Jitter is caused by the cumulative effects of queuing delays at the various transmission and switching points along the path of a given session through a network. Jitter will tend to have a maximum value for a given network path loosely defined by the congestion thresholds set for the nodes and trunks along that path. In general increasing Jitter will give way to packet loss as link load values exceed 50%-70% or node load values exceed 70%-90%, depending on the traffic profiles for the network (Rahalho, 1999).

In any case jitter buffers are required whenever real-time traffic egresses a packet network in order that the potential for under-run conditions is minimized. In a voice application “underrun” describes the situation where analogue output is starved of digital input and “clipping” occurs. These jitter buffers contribute to the end-to-end delay for the call.

Network routers help reduce jitter by flushing their transmission buffers before significant queues accumulate. And, since interframe times are typically 30 to 100 milliseconds on a single VoIP session, it is highly two consecutive packets from a single session will end up simultaneously in the same transmission queue. In well-engineered networks enough packets will usually be transmitted for maintenance of reasonable voice quality even during normal transient congestion conditions. Jitter buffer values suitable for the current public Internet should be expected to exceed 200 milliseconds. While this still allows the provisioning of services with reasonable quality, IPV6 and Winsock 2 will also help to reduce the required jitter buffer values to around 100 milliseconds.

IP Telephone Growth Factor

The rapid growth is being driven by a powerful combination of factors. Currently, IP telephony gives new operators an easy and cost-efficient way to compete with incumbent companies by undercutting their pricing regimes, while avoiding many of the regulatory barriers to standard voice provision. However, unlike other low-price telecom services such as callback, the success of IP telephony is not merely dependent on exploiting the artificial price differentials which will fade away once the prices of standard services are set on a more rational economic basis. It also offers a platform for the integration of voice and data communications for easier development of value-added services such as video conferencing, unified messaging and increased call center functionality. More generally, the switch to IP as the main delivery mechanism for telecom services in the future is looking increasingly likely. The most recent harbingers of this change are the new entrant operators such as Qwest and Level3. Their extensive deployment of high-bandwidth IP networks is likely to encourage a wholesale switch to IP, as the economies of scale they attain drive down IP provisioning costs even further and force traditional operators to abandon their legacy system. The technological goal of the IP telephony industry is the establishment of a single IP pipeline carrying all of a corporate's voice, video, and data traffic. The benefits are numerous and compelling. First, consolidation will bring both cost savings and possible simplified ease of use via centralized network management software (Essex, '99)

Conclusion

The telecommunications industry is changing at an astounding rate. IP telephony offers potential to provide a richer service experience through the use of Internet-based server technology and the combination of video and data services with basic voice services. The use of the Internet as a distribution channel, coupled with the low start-up costs will dramatically alter the voice service rate. The H.323 is comprehensive, yet flexible, and can be applied to voice and full multimedia video conferencing stations. H.323 applications are set to grow into the main stream market.

REFERENCES

1. Armstrong, Larry, Sanders. You are coming over loud and almost clear: Business Week, Oct. 27, No. 1, Vol.14, P.116
2. Business Week. Networks that does the new tricks: April 6, 1998, No. 3572, PP. 88-91.
3. Couch, Chris: Interoperability: The reason for Standard in VoIP. Western Communications Forum, Feb. 18, 1999.
4. DataBeam Corporation. A Primer on the H.323 Series Standard. Lexington, Kentucky, May 15, 1998.
5. Essex, David. IP Telephony: What does the future hold? Network Magazine, Jan. 1999, No.1, Vol.14, PP. 40-44.
6. Fuller, David. IP Telephony: Business Communications View of Voice-over-IP. Western Communications Forum, Feb. 17, 1999.
7. Halsall, Fred. Data Communications, Computer Networks and Open Systems, University of Wales, Swansea, Addison-Wesley Publishing Co. 1992.
8. Hughes, Joel. Quality of Service Challenges in IP Telephony, Western Communications Forum, Feb. 18, 1999.
9. Lakelin, Philip. IP Telephony: Opportunity or Treat, HYPERLINK http://www.iptelephony.org
10. Kiersted, Thomas. IP Telephony Vs PSTN: End user Satisfaction and Drivers, Western Communications Forum, Feb. 17, 1999.

11. Ong, Lyndon. Protocol Activities for IP, Western Communications Forum, Feb. 18, 1999.
12. Shay, Williams A. Understanding Data Communications and Networks. Cole Publishing Co., 1999.
13. Siemens. Internet Telephony Tutorial. http://www.iptelephony.org Walters, Rob. Computer Telephony Integration: Boston, Artech House, 1993